DOCUMENT RESUME

ED 395 770 SE 058 162

AUTHOR	Kumar, David
TITLE	Computers and Assessment in Science Education. ERIC Digest.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-SE-96-3
PUB DATE	Mar 96
CONTRACT	RI88062006
NOTE	4p.
AVAILABLE FROM	ERIC/CSMEE, 1929 Kenny Road, Columbus, OH 43210-1080.
PUB TYPE	Information Analyses - ERIC Clearinghouse Products (071)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Computer Simulation; *Computer Uses in Education; Elementary Secondary Education; *Science Instruction; *Student Evaluation
IDENTIFIERS	*Alternative Assessment; ERIC Digests

ABSTRACT

This digest discusses types of computer applications used in science assessment; underlying assumptions that guide their roles in assessment; and issues which relate to validity, gender equity, instructional delivery, mode of user interface, and responsibility to the public. Assessment applications for computers can be broadly classified into two categories: traditional and contemporary. Considering the need for aligning testing with education reform efforts underway, the focus of researchers should be the contemporary rather than traditional approaches to computer-based assessment. Among contemporary computer applications, simulations appear to hold promises for large-scale assessment. Several examples of simulations are given. Contains 17 references. (MKR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 395 770

Computers and Assessment in Science Education

by

David Kumar

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Computers and Assessment in Science Education

David Kumar

March 1996

ERIC Clearinghouse for Science, Mathematics, and Environmental Education DIGEST

Research and development efforts in alternative forms of assessment are on the rise in science education, along with growing interest in using computers in science assessment. A review of technological applications in science assessment was provided by Helgeson and Kumar (1993), recent developments in computer scorable, large scale tests were reported by Martinez and Bennett (1992), and a theme issue on computer-based science assessment was prepared by the *Journal of Science Education and Technology* (Vol. 4, Issue 1, 1995).

Types of computer applications in science assessment

Assessment applications for computers can be broadly classified into two categories: traditional and contemporary. In traditional applications the infrastructure is rigidly algorithmic. Examples include forced-choice and multiple-choice testing, grading, and record keeping. In more contemporary applications the infrastructure is quasi-algorithmic or non-linear in nature. Examples include constructed response testing, adaptive testing, figural response testing, simulations, and solution pathway analysis.

Forced-choice and multiple-choice testing are the most traditional applications of computers in testing and are often used to test low level knowledge acquisition. Considering the need for aligning testing with education reform efforts underway, the focus of researchers should be the contemporary rather than traditional approaches to computer-based assessment. Contemporary computer applications would allow for better analysis of the new kinds of process-oriented science instruction suggested by the science education reform efforts.

Among contemporary computer applications, simulations appear to hold promises for large scale assessment. A comparative study of computer simulations and hands-on tasks such as the classic "batteries and bulbs" activities shows that student outcomes are not much different between computer-based and hands-on assessments (Shavelson, Baxter, & Pine, 1992). Gorrell (1992) reported a computer simulation-based assessment for collecting process information in learning among undergraduates in behavior analysis.

In figural response testing students manipulate pictorial tasks on a computer screen using a mouse to solve problems. According to Martinez (1993) this approach has been found suitable for science assessment involving extensive graphics in disciplines such as stereochemistry and molecular biology.

Constructed response testing gives students the option of presenting their answers or solutions to a problem on a computerized grid sheet (Martinez & Bennett, 1992). The computer also grades student responses with a preset standard deviation range, thereby reducing the chances of students losing all credit for a question because they picked the wrong answer from a multiple response. Students receive partial credit for partially correct answers. Braun, Bennett, Frye and Soloway (1990) reported using expert systems for scoring constructed responses of high school students in Advanced Placement computer science courses. They found the expert system was able to score between 82% and 95% of the responses successfully and show a high correlation with a human grader on correctness.

In computerized adaptive testing the computer tailors a test according to an examinee's level of achievement and ability. For example, based upon the kind of response made to a question on a particular topic, the computer will decide whether to stay on the same topic to ask another question that will help the student review or clarify background knowledge, or proceed to a higher level question and to a different topic (Welch & Frick, 1993).

Using computers for analyzing solution pathways is another emerging trend. Gong, Venezky and Mioduser (1992) used a computer-based learning progress map incorporating a database for biology testing leading to valuable and interesting analyses of student performance. Young (1993) also reported an anchored assessment approach involving a videodisc anchor and interacting computer software. The software functions as a ledger for recording the kinds of information the student searches and uses for problem solving using the videodisc. According to Young this anchored assessment technique has provided information on student performance in the solution space which otherwise would be obtainable only through verbal protocols and extensive transcription of data. In another study a Pen-Point computer was used to study solution pathway procedures in solving classical multiple step modularity problems (Kumar & Helgeson, 1995). In this study, problem solving protocols were registered in a way similar to recording verbal protocols in think-aloud procedures.

Assumptions

The computer applications described here have several underlying assumptions that guide their roles in assessment. First, computers are information management tools and they provide enormous opportunities for gathering and managing a variety of assessment data. Second, computers provide a less obtrusive medium for students to express their thought processes in a problem solving task (Schneiderman, 1987). Third, computers function as an extended working memory thereby reducing the cognitive load on the problem solver (Rowe, 1993) and adding considerably to "logical/mathematical intelligence" (Moursund, 1994). Fourth, hypermedia computer systems provide a flexible non-linear environment whereby the moves and steps a problem solver takes during an interaction with the computer could be recorded for assessment (Kumar, 1994). Fifth, human computer interaction is not just a mechanical relationship. It is mediated by a hypothetical interface, the "computer technology-cognitive psychology interface", which is a complex interaction between human cognition and the computer environment (Kumar, Helgeson, & White, 1994). More research is needed to understand this interface that enables thinking and expression of thoughts while interacting with computers.

Issues

While computer applications and the underlying assumptions of the role of computers in assessment open up doors of opportunity for the development of innovative computer-based tools, they also raise serious issues. Some of the key issues relate to validity, gender equity, instructional deliv-

ERIC/CSMEE • 1929 Kenny Road • Columbus, Ohio 43210-1080

ery, the mode of user interface, and responsibility to the public.

Developing computer-based assessment tasks in science that are valid for large groups is an issue that has yet to be fully resolved (Wainer, 1993; Welch & Frick, 1993; Martinez, 1993; Shavelson et al., 1992; Kumar, 1994). Establishing the validity of computer tests by comparing them against traditional pencil-and-paper tests is a double-edged sword. The main purpose for bringing computers into assessment is to develop tasks that will provide more information about thinking processes that are difficult to obtain through standardized multiple-choice tests. On the other hand, for public accountability, it is very difficult to compare and justify computer-based assessment results that are aimed at individual or small groups against large scale traditional standardized tests. Also, public pressure to disclose the questions and answers of large scale examinations, such as the computerized adaptive tests, has raised concerns about maintaining test validity over a large period of time (Jacobson, 1994). Therefore, more research is needed in the development and evaluation of computer-based assessment applications that are valid on a large scale.

Computer-based learning has been shown to raise gender equity issues (Rowe, 1993). Considering testing as a part of learning in science, gender equity is a serious issue in computer-based testing. Assessment tasks that are more appealing to female students ought to be taken into account in developing computer-based tests in science. If computers are to be used for science assessment they must also be used as an integral part of science teaching and learning. Research and development efforts should focus on computer technology applications and instruction as well as assessment in science.

Using computers for performance assessment is an interactive experience, and it depends upon the mode of user interface that links the person and the computer. User interface devices such as keyboards, mice, light pens, and induction pens all have different effects on the performance of the individual at the computer (Shneiderman, 1987; Kumar & Helgeson, 1995). Due to progress in computer technology, using virtual reality to simulate hands-on assessment tasks may be useful for designing more effective computer-based performance assessment applications in terms of less obtrusive user interface and an increased sense of realism. Research efforts in computer-based assessment need to take into account the effects of user interfaces on student performance in order to develop more valid assessment tools using computers.

Considering the developments in contemporary approaches to computer-based science assessment, it is evident that computers are viable tools for assessment. However, the inherent issues discussed earlier should be addressed in order to make computer-based testing an acceptable practice for large scale assessment. Continued research and development efforts are needed in order to shape computers into tools that are unquestionably effective on a large scale for performance assessment in science education. Computer-based assessment remains a fertile field for research and development in science.

References

Braun, H.I., Bennett, R.E., Frye, D., & Soloway, E. (1990). Scoring constructed responses using expert systems. *Journal of Educational Measurement*, *27*(2), 93-108.

Gong, B., Venezky, R., & Mioduser, D. (1992). Instructional assessments: Lever for systemic change in science education classrooms. *Journal of Science Education and Technology*, *1*(3), 157-176.

Gorrell, J. (1992). Outcomes of using computer simulations. *Journal of Research on Computing in Education*, *24*(3), 359-356.

Helgeson, S.L. & Kumar, D.D. (1993). A review of educational technology in science assessment. *Journal of Computers in Mathematics and Science Teaching*, *12*(3/4), 227-243.

Jacobson, R.L. (1994). Computerized testing runs into trouble: Political and technical questions are raised. *The Chronicle of Higher Education*, *XL*(48), A16-A17.

Kumar, D.D., & Helgeson, S.L. (1995). Trends in computer applications in science assessment. *Journal of Science Education and Technology*, *4*(1), 29-36.

Kumar D.D., Helgeson, S.L. & White, A.L. (1994). Computer technology-cognitive psychology interface and science performance assessment. *Educational Technology Research and Development*, *42*(4), 6-16.

Kumar, D.D. (1994). Hypermedia: A Tool for alternative assessment? *Educational & Training Technology International*, *31*(1), 59-66.

Martinez, M.E., & Bennett, R.E. (1992). A review of automatically scorable constructed-response item types for large scale assessment. *Applied Measurement in Education*, *5*(2), 151-169.

Martinez, M.E. (1993). Item formats and mental abilities in biology assessment. *Journal of Computers in Mathematics and Science Teaching*, *12*(3/4), 289-301.

Moursund, D. (1994). Computers and human intelligence. *The Computing Teacher*, *21*(8), 5.

Rowe, H.A.H. (1993). *Learning with personal computers*. Victoria, Australia; Australian Council for Educational Research.

Shneiderman, B. (1987). *Designing the user interface*. New York: Addison Wesley.

Shavelson, R.J., Baxter, G.P. & Pine, J. (1992). Performance assessment: Political rhetoric and measurement reality. *Educational Researcher*, *21*(4), 22-27.

Wainer, H (1993). Measurement problems. *Journal of Educational Measurement*, *30*(1), 1021.

Welch, R.E. & Frick, T. (1993). Computerized adaptive testing in instructional settings. *Educational Technology Research & Development*, *41*(3), 47-62.

Young, M.F. (1993). Instructional design for situated learning. *Educational Technology Research and Development*, *41*(1), 40-50.

About the Author

Dr. David Kumar is an associate professor of science education at Florida Atlantic University. His areas of research include technology, assessment, evaluation, and policy.

SE 058 162

➤ This digest is in the public domain and may be freely reproduced. ➤

EDO-SE-96-3

This digest was funded by the Office of Educational Research and Improvement, U.S. Department of Education under contract no. RI-88062006. Opinions expressed in this digest do not necessarily reflect the positions or policies of OERI or the Department of Education.

OERI

ERIC

The Educational Resources Information Center is a nationwide information system initiated in 1966 by the U.S. Department of Education. ERIC has developed the largest and most frequently used education-related database in the world. For information, call 1-800-538-3742.